

Expurgation Exponent of Leaked Information in Privacy Amplification for Binary Sources

Shun Watanabe*

* Department of Information Science and Intelligent Systems, University of Tokushima, Tokushima, Japan,
Email: shun-wata@is.tokushima-u.ac.jp

Abstract—We investigate the privacy amplification problem in which Eve can observe the uniform binary source through a binary erasure channel (BEC) or a binary symmetric channel (BSC). For this problem, we derive the so-called expurgation exponent of the information leaked to Eve. The exponent is derived by relating the leaked information to the error probability of the linear code that is generated by the linear hash function used in the privacy amplification, which is also interesting in its own right. The derived exponent is larger than state-of-the-art exponent recently derived by Hayashi at low rate.

I. INTRODUCTION

In information theoretic key agreement problem [1], [2], [3], [4], [5], [6], legitimate parties need to distill a secret key from a random variable in the situation such that an eavesdropper can access to a random variable that is correlated to the legitimate parties' random variable. The privacy amplification is a technique to distill a secret key under the situation by using a (possibly random) function [7]. The security of distilled key is evaluated by various kinds of measures. In this paper, we focus on the leaked information, which is the mutual information between the distilled key and eavesdropper's random variable (the so-called strong security [8], [9]), because it is the strongest notion among security criterion [4] (see also [10, Appendix 3]).

The privacy amplification is usually conducted by using a family of universal 2 hash functions [11]. In [7], Bennett *et al.* evaluated ensemble averages of the leaked information for universal 2 families, and derived an upper bound on the leaked information by using the Rényi entropy of order 2. In [12], Renner and Wolf evaluated ensemble averages of the leaked information for universal 2 families, and derived an upper bound on the leaked information by using the smooth minimum entropy. In [10], Hayashi evaluated ensemble averages of the leaked information for universal 2 families, and derived a parametric upper bound on the leaked information by using the Rényi entropy of order $1 + \theta$. Concerning the exponential decreasing rate of the leaked information, the exponent derived by Hayashi's bound is state-of-the-art.

In noisy channel coding problem, the exponential decreasing rate of the error probability is also regarded as an important performance criterion of codes, and has been studied for a long time. The best exponent at high rates is the one derived by Gallager's random coding bound [13]. However, Gallager's exponent is not tight in general, and can be improved at low rates because the random code ensemble involves some bad

codes and those bad codes become dominant at low rates. The improved exponent by expurgating those bad codes is usually called the expurgation exponent [13], [14]. Similar improved exponents are also known in the context of the Slepian-Wolf coding [15], [16] or the quantum error correction [17].

The purpose of this paper is to show a security analog of above results, i.e., to derive an improved exponent of the leaked information in the privacy amplification at low rates. For this purpose, we concentrate our attention on the case such that the random variable possessed by the legitimate parties is the binary uniform source and the function used in the privacy amplification is a linear matrix.

We first consider the case such that the eavesdropper's random variable is generated via a binary erasure channel (BEC). For this case, we first relate the leaked information to the maximum likelihood (ML) decoding error probability of the linear code whose generator matrix is the one used in the privacy amplification. Then an improved exponent is derived by using the result of the expurgation exponent of linear codes.

It should be noted that a similar approach to relate the leaked information to the erasure error correction has been appeared in [18]. However in this paper, we directly relate the leaked information to the ML decoding error probability, which enables us to derive the improved exponent. It should be also noted that the approach in this paper is completely different from the error correction approach conventionally used to prove the so-called weak security and the problem pointed out in [19] does not apply to our approach.

Next, we consider the case such that the eavesdropper's random variable is generated via a binary symmetric channel (BSC). For this case, the technique used in the BEC case cannot be directly applied. Thus, we first reduce the BSC case to the BEC case by using the partial order between BSCs and BECs. The reduction turns out to be quite tight. Indeed, the exponent derived via this reduction is as good as Hayashi's exponent below the critical rate, and strictly better than Hayashi's exponent below the expurgation rate, which resemble the relation between the expurgation exponent and the random coding exponent of the noisy channel coding. Our results suggest that the privacy amplification with a universal 2 family is not necessarily optimal.

The rest of the paper is organized as follows. We first explain the problem formulation of the privacy amplification in Section II. Then, we consider the BEC case and the BSC case in Sections III and IV respectively. Conclusions are discussed

in Section V.

II. PROBLEM FORMULATION

Let (X^n, Z^n) be a correlated i.i.d. source with distribution P_{XZ} . The alphabet is denoted by $\mathcal{X} \times \mathcal{Z}$. In the privacy amplification problem, we are interested in generating the uniform random number on \mathcal{S}_n by using a function $f_n : \mathcal{X}^n \rightarrow \mathcal{S}_n$. The joint distribution of the generated random number and the side-information is given by

$$P_{S_n Z^n}(s_n, z^n) = \sum_{x^n \in f_n^{-1}(s_n)} P_{XZ}^n(x^n, z^n),$$

where $f_n^{-1}(s_n) = \{x^n \in \mathcal{X}^n : f_n(x^n) = s_n\}$.

The security is evaluated by the leaked information

$$I(f_n) = I(S_n; Z^n)$$

where $I(\cdot; \cdot)$ is the mutual information and we take the base of the logarithm to be e .

For given rate $R \geq 0$, we are interested in the exponential decreasing rate of $I(f_n)$, i.e.,

$$E(R; X|Z) = \sup \left\{ \liminf_{n \rightarrow \infty} -\frac{1}{n} \log I(f_n) : \liminf_{n \rightarrow \infty} \frac{1}{n} \log |\mathcal{S}_n| \geq R \right\}.$$

In the privacy amplification problem, we typically use the universal 2 hash family.

Definition 1: A family \mathcal{F}_n of functions $f_n : \mathcal{X}^n \rightarrow \mathcal{S}_n$ is called universal 2 if

$$\Pr\{F_n(x^n) = F_n(\hat{x}^n)\} \leq \frac{1}{|\mathcal{S}_n|}$$

for every $x^n \neq \hat{x}^n$, where F_n is the uniform random variable on \mathcal{F}_n .

For parameter θ , let

$$\begin{aligned} \psi(\theta; X|Z) &= -\log \sum_{x,z} P_{ZX}(x,z)^{1+\theta} P_Z(z)^{-\theta} \\ &= -\log \sum_{x,z} P_{XZ}(x,z) \exp[\theta \log P_{X|Z}(x|z)]. \end{aligned}$$

Hayashi derived the following lower bound on $E(R; X|Z)$.

Proposition 2 ([10]): For any universal 2 hash family \mathcal{F}_n , we have

$$\begin{aligned} E(R; X|Z) &\geq \liminf_{n \rightarrow \infty} -\frac{1}{n} \log \mathbb{E}_{\mathcal{F}_n}[I(f_n)] \\ &\geq E_r(R; X|Z) \\ &:= \max_{0 \leq \theta \leq 1} [\psi(\theta; X|Z) - \theta R], \end{aligned}$$

where $\mathbb{E}_{\mathcal{F}_n}$ means the average over randomly chosen function from \mathcal{F}_n .

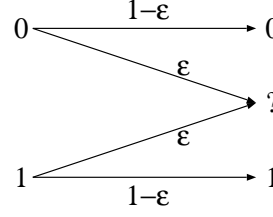


Fig. 1. The channel considered in Section III.

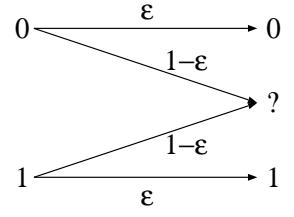


Fig. 2. The virtual channel considered in Section III.

III. SIDE-INFORMATION VIA BINARY ERASURE CHANNEL

In this section, we assume that X is the uniform binary source and Z is the output of the binary erasure channel (BEC) with erasure probability ε , i.e., $P_{XZ}(x, x) = \frac{1-\varepsilon}{2}$ and $P_{XZ}(x, ?) = \frac{\varepsilon}{2}$, where $?$ represent the erasure symbol (see Fig. 1). For given sequence z^n , let $\mathcal{J}(z^n) \subset \{1, \dots, n\}$ be the set of those indices such that $z_j = ?$. When the sequence z^n is obvious from the context, we abbreviate $\mathcal{J}(z^n)$ as \mathcal{J} .

In the rest of this paper, we concentrate on the linear function $f_n : \mathcal{X}^n \rightarrow \mathcal{S}_n$. Thus, we implicitly assume that $\mathcal{X} = \mathbb{F}_2$ and $\mathcal{S}_n = \mathbb{F}_2^k$ for some k , where \mathbb{F}_2 is the field of order 2. Let M_n be $k \times n$ matrix with entries in \mathbb{F}_2 . We consider function $f_n : x^n \rightarrow x^n M_n^T$ and the security criterion is denoted by $I(M_n)$. The sequence $x_{\mathcal{J}}^n$ is a subsequence of x^n that consist of the indices in \mathcal{J} , and the matrix $M_{\mathcal{J}}$ is a sub-matrix of M_n that consist of the columns in \mathcal{J} .

The following lemma was presented by Ozarow and Wyner.

Lemma 3 ([20]): We have

$$H(S_n | Z^n = z^n) \geq \text{rank}(M_{\mathcal{J}(z^n)})$$

for every z^n .

We consider the virtual BEC with erasure probability $1-\varepsilon$ (see Fig. 2), i.e., $P_{Y|X}(x|x) = \varepsilon$ and $P_{Y|X}(?|x) = 1-\varepsilon$. From Lemma 3, we have the following.

Theorem 4: Let C_n be the linear code whose generator matrix is M_n , and let $P_{ML}(C_n, 1-\varepsilon)$ be the maximum likelihood decoding error probability¹ of the code C_n over the BEC($1-\varepsilon$). Then, we have

$$I(M_n) \leq n P_{ML}(C_n, 1-\varepsilon).$$

Proof: Let $m^k \in \mathbb{F}_2^k$ is a message to be sent, and the encoded message $m^k M_n$ is sent over the BEC($1-\varepsilon$). Suppose that the received signal is y^n . If $\text{rank}(M_{\mathcal{J}(y^n)^c}) = k$, then the ML decoder output m^k , where $\mathcal{J}(y^n)^c = \{1, \dots, n\} \setminus \mathcal{J}(y^n)$ is the non erased bits. On the other hand, if $\text{rank}(M_{\mathcal{J}(y^n)^c}) < k$, there are plural messages that are compatible with y^n , and thus the ML decoder fail to output m^k . Therefore, the ML decoding error probability can be written as

$$\begin{aligned} P_{ML}(C_n, 1-\varepsilon) &= \sum_{\mathcal{J}^c \subset \{1, \dots, n\}} (1-\varepsilon)^{n-|\mathcal{J}^c|} \varepsilon^{|\mathcal{J}^c|} \mathbf{1}[\text{rank}(M_{\mathcal{J}^c}) < k]. \end{aligned}$$

¹Ties are counted as errors.

On the other hand, by using Lemma 3 and by noting that $H(S_n) \leq n$, we have

$$I(M_n) \leq n \sum_{\mathcal{J} \subset \{1, \dots, n\}} (1 - \varepsilon)^{n - |\mathcal{J}|} \varepsilon^{|\mathcal{J}|} \mathbf{1}[\text{rank}(M_{\mathcal{J}}) < k].$$

Thus, we have the assertion of the theorem. \blacksquare

By using a linear code achieving the Gilbert-Varshamov bound, we have the following.

Corollary 5: There exists a linear function $f_n : x^n \rightarrow x^n M_n^T$ such that

$$E(R; X|Z) \geq \liminf_{n \rightarrow \infty} -\frac{1}{n} \log I(f_n) \quad (1)$$

$$\geq \liminf_{n \rightarrow \infty} -\frac{1}{n} \log P_{ML}(\mathcal{C}_n, 1 - \varepsilon) \quad (2)$$

$$\geq E_x(R, 1 - \varepsilon) \quad (3)$$

$$:= \max_{\theta \geq 1} \left[\theta \{ \log 2 - R - \log(1 + (1 - \varepsilon)^{1/\theta}) \} \right]. \quad (4)$$

Proof: First note that the error probability of the channel coding and that of Slepian-Wolf coding (with full side-information) are the same for linear code and BEC. Thus, Csiszár's linear Slepian-Wolf code result [16] implies that there exists a code satisfying

$$\begin{aligned} & \liminf_{n \rightarrow \infty} -\frac{1}{n} \log P_{ML}(\mathcal{C}_n, 1 - \varepsilon) \\ & \geq \min_{H(W) \geq \log 2 - R} \left[(\log 2 - R) - H(W) + \right. \\ & \quad \left. \mathbb{E} \left[-\log \sum_{x,y} \sqrt{P_{XY}(x,y) P_{XY}(x+W,y)} \right] \right] \\ & = \min_{h(p) \geq \log 2 - R} [-p \log(1 - \varepsilon) + (\log 2 - R) - h(p)], \end{aligned} \quad (5)$$

where we set $P_W(1) = p$. Since the objective function of Eq. (5) is convex, by introducing

$$L(\lambda) := \min_p [-p \log(1 - \varepsilon) + (1 + \lambda)(\log 2 - R - h(p))]$$

for $\lambda \geq 0$, Eq. (5) can be written [21] as

$$\max_{\lambda \geq 0} L(\lambda).$$

By changing the variable as $\theta = 1 + \lambda$, Eq. (5) can be also written as

$$\max_{\theta \geq 1} L(\theta - 1) = \max_{\theta \geq 1} \left[\theta \{ \log 2 - R - \log(1 + (1 - \varepsilon)^{1/\theta}) \} \right].$$

Note that $E_x(R, 1 - \varepsilon)$ is the expurgation exponent for BEC($1 - \varepsilon$) [22].

Remark 6: It should be noted that

$$E_r(R; X|Z) = E_r(R, 1 - \varepsilon) \quad (6)$$

$$:= \max_{0 \leq \theta \leq 1} \left[-\log \left\{ (1 - \varepsilon) + \frac{1}{2^\theta} \varepsilon \right\} - \theta R \right]. \quad (7)$$

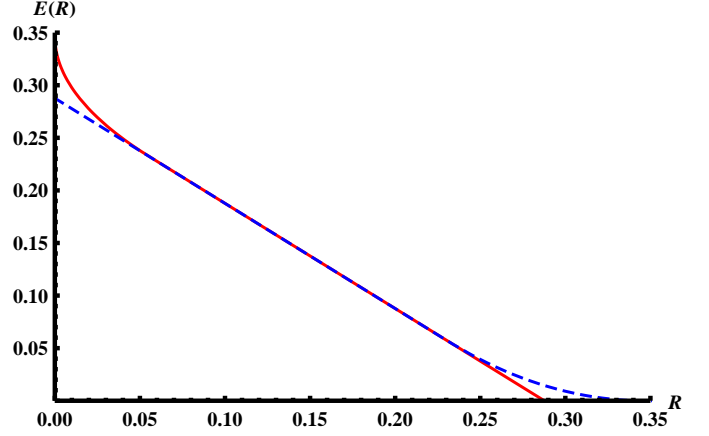


Fig. 3. Comparison of $E_r(R, 1 - \varepsilon)$ (dashed line) and $E_x(R, 1 - \varepsilon)$ (solid line) for $\varepsilon = 0.5$.

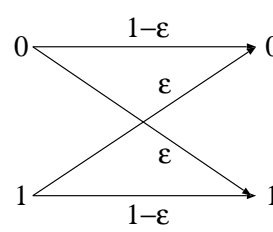


Fig. 4. The channel considered in Section IV.

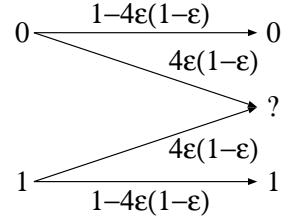


Fig. 5. The virtual channel considered in Section IV. This channel is less noisy than the BSC in Fig. 4.

Since $E_r(R, 1 - \varepsilon)$ is the random coding exponent for BEC($1 - \varepsilon$) [22], Hayashi's exponent can be also derived from Theorem 4.

From Eq. (3) and Eq. (6) and known facts on the exponents, we find that the exponent of PA in Corollary 5 is larger than that in Proposition 2 for low R . These exponents are compared in Fig. 3 for $\varepsilon = 0.5$. We find that $E_x(R, 1 - \varepsilon)$ is strictly larger than $E_r(R, 1 - \varepsilon)$ at low rates.

IV. SIDE-INFORMATION VIA BINARY SYMMETRIC CHANNEL

In this section, we assume that X is the uniform binary source and Z is the output of the binary symmetric channel (BSC) with crossover probability ε , i.e., $P_{XZ}(x, x) = \frac{1 - \varepsilon}{2}$ and $P_{XZ}(x, x + 1) = \frac{\varepsilon}{2}$ (see Fig. 4). Let \bar{Z} be the output of BEC($4\varepsilon(1 - \varepsilon)$) with input X . Since BEC($4\varepsilon(1 - \varepsilon)$) (see Fig. 5) is less noisy than BSC(ε) [23], we have

$$I(S_n; Z^n) \leq I(S_n; \bar{Z}^n).$$

Thus, Corollary 5 can be applied to the case considered in this section.

Theorem 7: Let \bar{Z} be the output of BEC($4\varepsilon(1 - \varepsilon)$) with input X . Then, we have

$$\begin{aligned} E(R; X|Z) & \geq E(R; X|\bar{Z}) \\ & \geq E_x(R, 1 - 4\varepsilon(1 - \varepsilon)). \end{aligned}$$

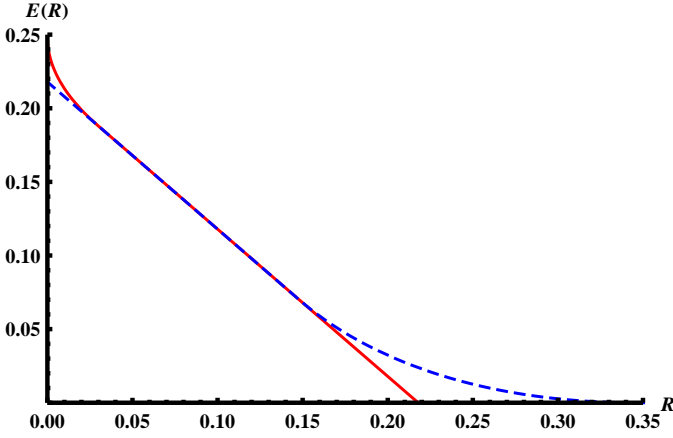


Fig. 6. Comparison of $E_r(R, X|Z)$ (dashed line) and $E_x(R, 1 - 4\varepsilon(1 - \varepsilon))$ (solid line) for BSC(0.11).

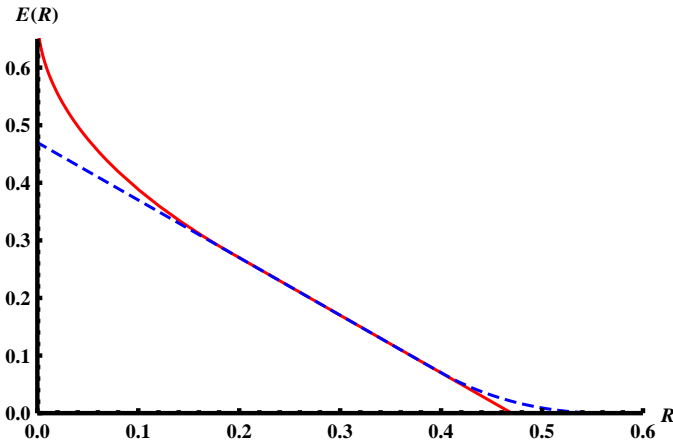


Fig. 7. Comparison of $E_r(R, X|Z)$ (dashed line) and $E_x(R, 1 - 4\varepsilon(1 - \varepsilon))$ (solid line) for BSC(0.25).

Hayashi's exponent for BSC(ε) is

$$E_r(R; X|Z) = \max_{0 \leq \theta \leq 1} [-\log \{(1 - \varepsilon)^{1+\theta} + \varepsilon^{1+\theta}\} - \theta R].$$

The exponents are compared in Fig. 6 and Fig. 7 for $\varepsilon = 0.11$ and 0.25 respectively.

Let $R_{cr}(\varepsilon)$ be the critical rate, i.e., the largest rate such that the optimization in $E_r(R; X|Z)$ is achieved by $\theta = 1$. Then, for $R \leq R_{cr}(\varepsilon)$, we have

$$E_r(R; X|Z) = -\log\{(1 - \varepsilon)^2 + \varepsilon^2\} - R.$$

On the other hand, let $R_x(\varepsilon)$ be the expurgation rate, i.e., the smallest rate such that the optimization in $E_x(R, 1 - 4\varepsilon(1 - \varepsilon))$ is achieved by $\theta = 1$. Then, for $R_x(\varepsilon) \leq R$, we have

$$\begin{aligned} E_x(R, 1 - 4\varepsilon(1 - \varepsilon)) \\ &= \log 2 - R - \log(1 + 1 - 4\varepsilon(1 - \varepsilon)) \\ &= -\log\{(1 - \varepsilon)^2 + \varepsilon^2\} - R. \end{aligned}$$

Thus, for $R_x(\varepsilon) \leq R \leq R_{cr}(\varepsilon)$, $E_r(R; X|Z) = E_x(R, 1 - 4\varepsilon(1 - \varepsilon))$, which can be also observed in Fig. 6 and Fig. 7.

We also find that $E_x(R, 1 - 4\varepsilon(1 - \varepsilon))$ is strictly larger than $E_r(R; X|Z)$ at low rates.

V. CONCLUSION

For the BEC case and the BSC case, we derived the expurgation exponent of the leaked information in the privacy amplification. The technique to relate the leaked information to the ML decoding error probability heavily relies on the specific structure of the BEC. Thus, to derive the expurgation exponent for general cases, a method to expurgate bad functions directly might be needed.

Hayashi derived a quantum counter part of Proposition 2 in [24]. It is also interesting to derive the expurgation exponent in the privacy amplification for quantum adversary. For the case such that the eavesdropper's information is generated via the complementary channel of a Pauli channel, the technique to relate the leaked information to the ML decoding error probability is already known [25]², and it is not difficult to derive the expurgation exponent. In general, more refined technique is needed. These topics will be investigated in elsewhere.

ACKNOWLEDGMENT

The author would like to thank Prof. Yasutada Oohama for valuable discussion. The author also would like to thank Prof. Prakash Narayan for inviting the author to the workshop. This research is partly supported by Grand-in-Aid for Young Scientist(B):2376033700 and Grand-in-Aid for Scientific Research(A):2324607101.

REFERENCES

- [1] U. Maurer, "Secret key agreement by public discussion from common information," *IEEE Trans. Inform. Theory*, vol. 39, no. 3, pp. 733–742, May 1993.
- [2] R. Ahlswede and I. Csiszár, "Common randomness in information theory and cryptography—part I: Secret sharing," *IEEE Trans. Inform. Theory*, vol. 39, no. 4, pp. 1121–1132, July 1993.
- [3] I. Csiszár and P. Narayan, "Common randomness and secret key generation with a helper," *IEEE Trans. Inform. Theory*, vol. 46, no. 2, pp. 344–366, March 2000.
- [4] —, "Secrecy capacities for multiple terminals," *IEEE Trans. Inform. Theory*, vol. 50, no. 12, pp. 3047–3061, December 2004.
- [5] —, "Secrecy capacities for multiterminal channel models," *IEEE Trans. Inform. Theory*, vol. 54, no. 6, pp. 2437–2452, June 2008.
- [6] A. A. Gohari and V. Anantharam, "Information-theoretic key agreement of multiple terminals—part I," *IEEE Trans. Inform. Theory*, vol. 56, no. 8, pp. 3973–3996, August 2010.
- [7] C. H. Bennett, G. Brassard, C. Crépeau, and U. Maurer, "Generalized privacy amplification," *IEEE Trans. Inform. Theory*, vol. 41, no. 6, pp. 1915–1923, Nov. 1995.
- [8] U. Maurer, "The strong secret key rate of discrete random triples," in *Communication and Cryptography—Two Sides of One Tapestry*. Kluwer Academic Publishers, 1994, pp. 271–285.
- [9] I. Csiszár, "Almost independence and secrecy capacity," *Problems of Information Transmission*, vol. 32, no. 1, pp. 40–47, 1996.
- [10] M. Hayashi, "Exponential decreasing rate of leaked information in universal random privacy amplification," *IEEE Trans. Inform. Theory*, vol. 57, no. 6, pp. 3989–4001, June 2011, arXiv:0904.0308.

²Although a result in the classical information theory is typically a commutative special case of the quantum counter part, this is not the case for the result shown in this paper. Indeed, the result in [25] is derived by using the relation between the eavesdropper's information gain and the amount of phase error caused in the main channel, which is a unique feature of quantum mechanics and there is no classical counter part.

- [11] J. L. Carter and M. N. Wegman, "Universal classes of hash functions," *Journal of Computer and System Sciences*, vol. 18, pp. 143–154, 1979.
- [12] R. Renner and S. Wolf, "Simple and tight bound for information reconciliation and privacy amplification," in *Advances in Cryptology – ASIACRYPT 2005*, ser. Lecture Notes in Computer Science, vol. 3788. Springer-Verlag, 2005, pp. 199–216.
- [13] R. G. Gallager, "A simple derivation of the coding theorem and some applications," *IEEE Trans. Inform. Theory*, vol. 11, no. 1, pp. 3–18, January 1965.
- [14] A. Barg and G. D. Forney, "Random codes: Minimum distances and error exponents," *IEEE Trans. Inform. Theory*, vol. 48, no. 9, pp. 2568–2573, September 2002.
- [15] I. Csiszár and Körner, "Graph decomposition: A new key to coding theorems," *IEEE Trans. Inform. Theory*, vol. 27, no. 1, pp. 5–12, January 1981.
- [16] I. Csiszár, "Linear codes for sources and source networks: Error exponents, universal coding," *IEEE Trans. Inform. Theory*, vol. 28, no. 4, pp. 585–592, July 1982.
- [17] A. Barg, "A low-rate bound on the reliability of a quantum discrete memoryless channel," *IEEE Trans. Inform. Theory*, vol. 48, no. 12, pp. 3096–3100, December 2002.
- [18] A. Thangaraj, S. Dihidar, A. R. Calderbank, S. W. McLaughlin, and J.-M. Merolla, "Application of ldpc codes to the wiretap channel," *IEEE Trans. Inform. Theory*, vol. 53, no. 8, pp. 2933–2945, August 2007.
- [19] S. Watanabe, R. Matsumoto, and T. Uyematsu, "Strongly secure privacy amplification cannot be obtained by encoder of slepian-wolf code," *IEEE Trans. Fundamentals*, vol. E93A, no. 9, pp. 1650–1659, September 2010, arXiv:0906.2582.
- [20] L. Ozarow and A. D. Wyner, "Wire-tap channel II," *AT&T Bell Labs. Technical Journal*, vol. 63, no. 10, pp. 2135–2157, December 1984.
- [21] S. Boyd and L. Vandenberghe, *Convex Optimization*. Cambridge University Press, 2004.
- [22] R. G. Gallager, *Information Theory and Reliable Communication*. John Wiley & Sons, 1968.
- [23] C. Nair, "Capacity regions of two new classes of two-receiver broadcast channels," *IEEE Trans. Inform. Theory*, vol. 56, no. 9, pp. 4207–4214, September 2010.
- [24] M. Hayashi, "Precise evaluation of leaked information with universal 2 privacy amplification in the presence of quantum attacker," arXiv:1202.0601.
- [25] —, "Practical evaluation of security for quantum key distribution," *Phys. Rev. A*, vol. 74, no. 2, p. 022307, August 2006, arXiv:quant-ph/0602113.